

Arquitectura de información para sitios de gran porte

Por Daniel Mordecki

Resumen: *Se intenta aportar un modelo de arquitectura de la información que sirva como base en la implementación de sitios web de organizaciones muy grandes y complejas, las cuales abarcan un universo de información amplio y variado.*

Palabras clave: *Arquitectura de la información, IA, Facetas, Navegación.*

Title: **Information architecture for large-sized sites**

Abstract: *We attempt to provide a model of information architecture that serves as a basis for the implementation of web sites of very large and complex organizations that publish a vast and diverse amount of information.*

Keywords: *Information architecture, IA, Facets, Navigation.*

Mordecki, Daniel. Arquitectura de información para sitios de gran porte. *El profesional de la información*, 2010, julio-agosto, v. 19, n. 4, pp. 368-373.

DOI: 10.3145/epi.2010.jul.05



Daniel Mordecki es director de Concreta (www.concreta.com.uy), empresa dedicada a la usabilidad y estrategia en internet. Cuenta con una vasta experiencia en dirección de proyectos tecnológicos, así como en diseño de la interacción y usabilidad, habiendo realizado estas tareas en numerosos proyectos. Es además docente en el área de e-business en la Universidad ORT y autor del libro "Pensar primero - Sepa por qué los programadores le contestan 'no se puede' cada vez que usted pide algo razonable y sencillo" (www.mordecki.com/libro). daniel@mordecki.com

1. Problemas que muestra la experiencia

Crear una arquitectura de la información (IA) para un universo de información extenso y variado se nos ha tornado en la práctica mucho más difícil y elusivo de lo que parecía al comienzo.

La organización de la información partiendo de la perspectiva de los visitantes, por lo menos en nuestra experiencia, nos ha planteado problemas en varios frentes. Abordarlos no resultó sencillo. Crear una solución efectiva ha implicado aplicar la teoría de una forma creativa y libre de apriorismos.

Como todos los que trabajamos en usabilidad y la constelación de disciplinas relacionadas, en *Concreta* participamos de esa persistente y empecinada batalla por reemplazar la perspectiva interna con la que se expone la información en numerosos sitios web por la perspectiva del usuario. El éxito en esta batalla depende en gran parte de la capacidad de crear una IA que soporte la organización de los contenidos y su navegación desde la visión del visitante (**Krug**, 2005).

En los diversos casos a los que nos hemos enfrentado hemos tenido suerte variada. A lo largo de ese camino, y en particular cuando se trata de organizaciones de gran porte, nos hemos encontrado con algunos problemas y algunas condiciones que al principio nos resultaron paradójicas y fuertemente contradictorias

con el marco teórico en el que nos basábamos. La organización de la información lisa y llana en función de la perspectiva de los visitantes resultaba fácil al comienzo y para los contenidos de alto tráfico, pero presentaba una serie de problemas muy difíciles de sortear al intentar abarcar el total de la información (**Morville** y **Rosenfeld**, 2006).

Tal vez sea difícil transmitir la sensación de frustración que esta situación representa: entender claramente qué es lo que hay que hacer y no poder proporcionar caminos sólidos y viables para hacerlo, no porque el cliente no quiera o pueda entenderlo, sino porque el marco teórico y nuestras propias limitaciones no nos permiten aportar soluciones realmente viables a los escollos a sortear. Describimos a continuación los más relevantes:

Grandes zonas grises en la clasificación de contenidos

La visión del usuario permite muchas veces ubicar rápidamente áreas de interés, temas o audiencias específicas que operan como primer nivel de la clasificación. Por ejemplo, en la seguridad social el perfil del visitante resulta un punto de partida adecuado: empleados, empresarios, jubilados, madres, niños, trabajadores independientes, etc.

Así podemos rápidamente encontrar para cada audiencia los derechos, las obligaciones y los servicios,

Artículo recibido el 25-04-10

Aceptación definitiva: 08-07-10

entre otros, que son en definitiva la información más relevante para cada visitante que se identifique con uno de los perfiles definidos. Pero ¿dónde colocar la normativa? Porque las leyes no están separadas por audiencia: los derechos de los empleados son obligaciones de los empresarios, que se transformarán con el tiempo en derechos de los jubilados. Parecido para la cobertura de salud, que cubre a empleados, algunos empresarios, algunos jubilados, trabajadores independientes, madres y niños por igual.

Claro que siempre surgen soluciones, atajos y alternativas: duplicar la información, asignarla a una audiencia y redirigir a los visitantes desde las otras audiencias, juntar dos categorías o crear una nueva y muchas otras más. Pero cuando se mira el resultado final, la clasificación de todo el universo de información, se perciben grandes áreas clasificadas “a la fuerza” y no en la forma natural que debería ser el resultado de una IA sólida.

Los criterios de clasificación tienen en general zonas grises y problemas de ambigüedad, pero cuando una clasificación es sólida estas zonas grises no deberían ser significativas. Para nuestro equipo de trabajo, es un requisito que sean menores a un 15 ó 20% del total de contenidos al momento de implantarla. Nuestra experiencia es que en muchos casos nos enfrentábamos a que tal vez la mitad de los documentos resultaban muy difíciles de clasificar.

Bolsones de “otros”

La categoría “otros”, que podríamos traducir como “pusimos acá lo que no supimos donde ubicar”, es el último recurso para incluir en alguna opción de menú el contenido que no calza en ningún lugar de la taxonomía. A veces tenemos suerte y se nos ocurre una etiqueta más útil y otras veces algún pícaro coloca un “Información de interés” para dar acceso a una colección ecléctica de formularios, contenidos e información que no queda bien en ningún otro lado.

Cuando la clasificación es acertada, el uso de este recurso es marginal y tiene un volumen muy pequeño. En nuestra experiencia, cuando adoptábamos al inicio la clasificación aparentemente más natural, nos enfrentamos a la alternativa de asumir grandes bolsones de contenidos en las opciones del tipo “otros”. Cuando se clasificaban meticulosamente uno a uno los contenidos, comenzaban a acumularse grupos de páginas e informaciones que ni siquiera podían ser ubicadas en la “menos mala” de las opciones, y se tornaban en populosos candidatos a la categoría otros.

Responsabilidad por la actualización

La diferencia radical entre la visión del usuario y la visión interna de una gran organización tiene una

consecuencia adicional a la hora de construir la IA. Si tomamos la primera, numerosas páginas estarán compuestas de piezas de información cuya responsabilidad recae en áreas distintas del organigrama institucional. Algo que es natural y razonable para la página principal y para algunas sub-homes, se transforma en un verdadero problema organizativo y de competencias cuando se multiplica significativamente.

El estilo de redacción, la frecuencia de actualización, la forma de interactuar con el visitante y muchos otros aspectos se tornan un escollo cuando las páginas son un rompecabezas con múltiples responsables. ¿Cómo sintonizar un párrafo creado por el área de Marketing con uno creado por el área Legal?

El proceso de crear la IA subyacente a un sitio no sólo tiene que permitir clasificar la información al momento de implementarla, sino que tiene que mantenerse razonablemente estable en el tiempo, hasta que los avatares de la vida exijan una nueva clasificación.

La incapacidad para delimitar los responsables de la información con claridad, así como la falta de claridad en los procesos de actualización han sido, otra vez en base a nuestra experiencia, un escollo casi imposible de saltar.

Ya sea porque la clasificación tal como estaba era detenida antes de ser presentada al cliente o porque el propio cliente se negaba a aceptarla, la mayoría de las veces era necesario volver atrás y realizar modificaciones para que hubiera una correspondencia más directa entre las áreas responsables por la información y las áreas del sitio determinadas por la IA subyacente, lo que determinaba invariablemente que el sitio se acercase más a la estructura interna de la organización a la vez que se aleja de la visión del usuario, algo que sólo puede ser valorado como una batalla perdida para un profesional de nuestra área.

Más de una visión

Si bien se trata de un problema menos radical que los anteriores, cuando el universo de información a categorizar es realmente muy amplio y variado, elegir una visión para definir la taxonomía es un problema. Volviendo al caso de la seguridad social, si se eligen los perfiles (empleados, empresarios, etc.) quedan relegadas las áreas: salud, vivienda, trabajo, etc. Y como éste numerosos ejemplos.

La IA debe tener la firmeza necesaria para organizar todos los documentos, pero ello no debe ser a costa de la flexibilidad. No sólo debe dar cabida a las visiones distintas que pudieran existir, sino que debe estar preparada para que dentro de ciertos parámetros estas visiones varíen con el tiempo. Cualquier hecho de magnitud probablemente cambie la relevancia relativa de las

Register for free at <https://www.scipedia.com> to download the version without the watermark

distintas partes de un sitio y es deseable que la IA provea herramientas para reflejar esta nueva visión de los contenidos y sus importancias relativas sin alteraciones sustanciales a la taxonomía de base. Por ejemplo, la promulgación de una ley de seguridad social no debería implicar el cambio del sitio actual por uno nuevo, sino una reorganización de la importancia de las distintas áreas, que además se irá modificando con el tiempo, en la medida que la aplicación de la ley avance.

Esto es algo que no se puede alcanzar si un cambio en la visión implica sustituir la clasificación principal en torno a la cual se construye la taxonomía, lo que implicará a su vez la reclasificación de la mayoría de los contenidos, que no es otra cosa que la sustitución de una IA por otra.

2. IA y navegación facetada

La solución a estos problemas surgió del trabajo incansable de los profesionales involucrados en la creación de las taxonomías y categorización de los documentos, que se negaban a admitir que no surgiera una propuesta capaz de mantener la perspectiva del usuario a la vez que resolvía los problemas a los que nos enfrentábamos. Y surgió a partir de la aplicación de una técnica que forma parte de la valija de herramientas de la IA desde tiempo atrás, utilizándola con un enfoque distinto al habitual: las facetas.

Según Wikipedia “Una clasificación facetada permite clasificar de múltiples formas a cada objeto, permitiendo ordenar las clasificaciones de formas múltiples, en vez del orden único y pre-determinado de una taxonomía. Una faceta incluye ‘*características definidas claramente, mutuamente excluyentes y que abarcan en forma colectiva todos los aspectos y propiedades de una clase o tema específico*’¹. Por ejemplo, una colección de libros podrá ser clasificada utilizando una faceta por autor, una por tema, una por fecha, etc.”².

La literatura de IA coincide en citar como origen de la clasificación facetada una propuesta denominada “Colon classification” introducida por el bibliotecólogo y matemático indio **Shiyali R. Ranganathan**³. Actualmente la utilización de facetas está muy vinculada a los sistemas de búsqueda y técnicas de filtrado en base a complejos sistemas de metadatos.

Taxonomía básica

Tal vez el elemento menos ortodoxo de la propuesta fue aceptar que la taxonomía básica estuviera apoyada en la estructura interna de la organización. Las categorías fueron repensadas para garantizar una granularidad, completitud y equilibrio adecuados, pero el resultado no se distanciaba demasiado del organigrama.

Si bien aparentemente esta solución nos alejaba del objetivo de acercar la organización de la información al usuario, algo que resolverá más adelante la aplicación de facetas, aportaba soluciones a los problemas a los que nos enfrentábamos:

Completitud: todas las piezas de información del sitio tienen un dueño, y en prácticamente todos los casos el dueño es único y fácil de identificar en la organización. Eso permite rápidamente clasificar todos los documentos del universo de información, garantizando que hay en la taxonomía un camino de navegación a todos y cada uno de ellos.

Responsabilidad: es el dueño de la información quien determina cómo y cuándo actualizar la información. En una gran organización es muy difícil que un área ceda esta responsabilidad, por lo que acercar la taxonomía a la estructura organizacional elimina esta barrera.

Exclusión mutua: si bien permanecen áreas de superposición, la mayor parte de los contenidos tiene origen y son propiedad de una única área de la organización y por tanto serán parte de una única categoría de la taxonomía, eliminando las zonas grises, las necesidades de duplicación y las ubicaciones ambiguas.

Navegación facetada

Por sobre la taxonomía, y fuertemente apoyada en ésta, se construyen nuevas categorizaciones que responden a las distintas visiones o facetas de la información. El hecho de que exista una taxonomía base completa es un elemento sustancial, ya que releva a estas clasificaciones de algunas exigencias que de otra forma le serían obligatorias (Segaran, 2007).

La representación en el sitio web de cada faceta es un minisitio, muy pequeño, que reúne la información relevante sobre un área de interés en la perspectiva de un perfil de usuario o área temática. El minisitio, más allá de su apariencia y de la enorme utilidad que tiene, no es un repositorio de contenido sino fundamentalmente una herramienta de navegación. Un conjunto coherente y bien organizado de resúmenes y vínculos relevantes sobre un tema, complementado con un pequeño conjunto de contenidos sólo cuando es estrictamente necesario.

La implementación de las facetas tiene ventajas y propiedades que resuelven las carencias de la taxonomía base. En primer lugar, las clasificaciones facetadas pueden incluir libremente los documentos que resulten convenientes, sin necesidad de ser completas, priorizando el foco frente a la cobertura. Esto aporta integridad y coherencia a la información de cada faceta, juntando en un solo lugar la información valiosa para una audiencia, para un tema o para una perspectiva del problema.

A esto se suma en segundo lugar que las clasificaciones facetadas incluyen los documentos, pero no

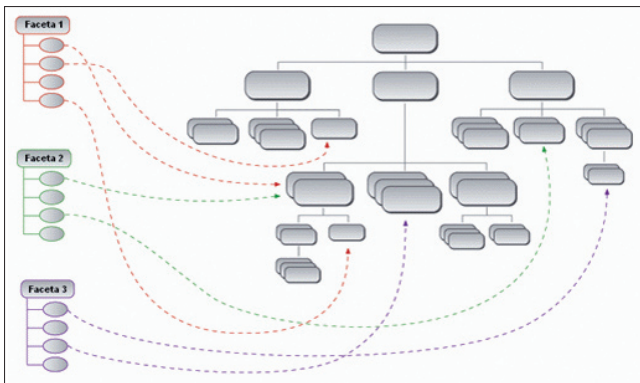


Figura 1. IA base y minisitios que implementan las facetas

tienen responsabilidad por el detalle del contenido ni por su actualización. Funcionan como agrupaciones de información relevante en la perspectiva del visitante, pero no son más que un esqueleto que organiza la información. La categorización facetada indica que para el tema X los documentos relevantes son el A, el B y el C, pero cada uno de dichos documentos sigue perteneciendo y siendo actualizado por el área de la organización que tienen la responsabilidad natural de hacerlo.

Por último y no por ello menos importante, se pueden crear cuantas clasificaciones facetadas sea necesario. En los niveles más altos del sistema de navegación se abren los abanicos de opciones facetadas. Los tests muestran que los usuarios se identifican rápidamente con el perfil o la visión que mejor aplica a sus necesidades de información de ese momento particular, y la facetación los conduce rápidamente al minisitio que contiene el acceso a la información que buscan, un micromundo donde mágicamente desapareció todo lo que no es relevante.

Importancia relativa de cada elemento en la pantalla

El elemento decisivo en la aplicación de la IA propuesta es que la ubicación principal en la pantalla la ocupa la navegación facetada y no la taxonomía principal. Que la taxonomía principal, reflejada en lo que podemos llamar “menú principal”, no ocupe el lugar más destacado de la navegación es también una decisión poco ortodoxa, pero nos permitía volver al origen del problema y aportar la solución: cuando el visitante arriba al sitio, el lugar más destacado lo ocupa la navegación facetada, es decir, la agrupación de contenidos por audiencias o temas de interés desde la perspectiva del usuario.

En el diseño de las herramientas de navegación y principalmente de la página principal del sitio web se debe poner especial cuidado en que la navegación por facetas provea a los públicos primarios un rápido camino, si fuera posible de un solo clic, hacia el dominio de información que el visitante necesita (Smiley y Pugh, 2009).

Las organizaciones de gran porte tienen además muchos públicos secundarios. Casos típicos son la prensa, el gobierno, los extranjeros entre muchos otros. Los públicos secundarios deberán recurrir a herramientas de navegación que tienen una ubicación y un destaque también secundario, y esto es acertado ya que responde precisamente a las causas que hicieron que la organización no coloque a estos públicos en los lugares de privilegio.

El menú principal, que refleja el primer nivel de la taxonomía base, se constituye en un ancla que permite a cualquier visitante navegar hacia cualquier contenido, inclusive los de tráfico casi insignificante, aunque ello implique en algunos casos pasar por tres o cuatro niveles de la taxonomía. Si además se cuenta con un buen sistema de búsqueda (Grossman; Frieder, 2004; Hearst, 2009) el acceso a los contenidos de menos tráfico queda asegurado.

3. Algunos ejemplos

Instituto de Seguridad Social de Uruguay

Tal como se indicó en los ejemplos a lo largo del artículo, cualquier organismo estatal que tenga como cometido la seguridad social tendrá en su público objetivo prácticamente a toda la población del país: los

Register for free at <https://www.scipedia.com> to download the version without the watermark



Figura 2. Instituto de Seguridad Social de Uruguay, <http://www.bps.gub.uy/>

Información sobre:	Información para:	Madres y niños	Empresas	Trabajadores
Procesos de reforma	Jubilados y pensionistas			
- Plan de equidad	- Derechos y obligaciones	- Derechos y obligaciones	- Derechos y Obligaciones	- Derechos y obligaciones
- Reforma tributaria	- Calendario de pagos	- Calendario de pagos	- Calendario de vencimientos	- Calendario de pagos
- Reforma de salud	- Servicios en Línea	- Servicios en Línea	- Servicios en Línea	- Servicios en Línea

Figura 3. Pie de la web del Instituto de Seguridad Social de Uruguay
<http://www.bps.gub.uy/>

trabajadores activos, los empresarios, los niños y los jubilados. A eso se suman otros enfoques temáticos y otras clasificaciones de públicos en función de visiones complementarias a la principal. A ese problema se enfrenta también el *Instituto de Previsión Social de Uruguay*⁴.

En la página principal, podemos encontrar los elementos típicos de una IA como la propuesta:

Menú facetado: ocupando un lugar de relevancia se encuentra el menú de acceso para las audiencias principales. El espacio que ocupa es tan grande que ni siquiera tiene apariencia de menú, pero en el fondo no es otra cosa que el acceso a los minisitios que agrupan y ordenan la información más importante para las cuatro audiencias destacadas del instituto.

Menú principal: está constituido por el primer nivel de la taxonomía base y garantiza un camino de navegación a cualquier contenido del sitio. Está integrado al cabezal y se repite en todas las páginas, de modo de constituirse en un apoyo

El pie: repite en todas las páginas las opciones de navegación por facetas (información para) y agrega una facetación de menos relevancia que la principal, bajo el título “Información sobre” que utiliza como criterio los procesos de reforma que afectan a la seguridad social. Si bien en la página principal el pie puede llegar a ser un elemento de relevancia menor, el hecho de que se repita incambiado en todas las páginas sin duda eleva significativamente su importancia.

IBM

En el sitio corporativo de IBM también se puede detectar una IA facetada, similar a la propuesta más arriba.

Menú principal: es el clásico menú de un proveedor de productos

y servicios informáticos. Responde fuertemente a la estructura de la empresa y en general hay que llegar al tercer nivel para encontrar los primeros contenidos “útiles” para un usuario, pero garantiza un camino de navegación a cualquier contenido del sitio.

Menú facetado: debajo del rotativo de novedades se despliega un gran menú que divide el contenido según 6 criterios, generando un gran espacio de facetas en las que se presume que será muy fácil encontrar lo que el visitante está buscando.

Localización: es muy interesante apreciar como en un sitio local, en este caso de Uruguay, el menú principal es una traducción literal del sitio corporativo mientras que el facetado es una personalización específica para los públicos que atiende el sitio. En este caso la facetación se realiza solamente por 3 criterios.



Figura 4. Home del sitio corporativo de IBM, <http://www.ibm.com>



Figura 5. Home del sitio de IBM en Uruguay, <http://www.ibm.com/uy>

Sony USA

Con un enfoque gráfico completamente distinto y basado en una idea tecnológica también diametralmente opuesta a los dos anteriores, ya que el sitio es, por lo menos en su página principal, una única y monolítica película flash, el sitio de Sony USA deja ver que se apoya en una IA como la propuesta.

El menú principal, a pesar de la gráfica, el diseño, los colores y la película flash, parece calcado del sitio de IBM y de cualquier otro sitio corporativo de una empresa de tecnología.

El menú facetado en este caso tiene un único criterio, que segmenta a los usuarios por la actividad que desea realizar. A pesar de la apariencia, al hacer clic en cualquiera de los cuadros se despliega un minisitio totalmente integrado a la película flash que reúne las opciones relativas a la categoría del título.

4. Conclusiones

En resumen, cuando nos enfrentamos a un amplio universo de información el camino para acercar la organización de los contenidos es indirecto:

Register for free at <https://www.scipedia.com> to download the version without the watermark

Primero hay que construir una base sólida, apoyada en la estructura de la organización para garantizar completitud, evitar problemas de repetición, zonas grises y generar un proceso de actualización fluida.

Luego se generan facetas según perfiles o áreas de interés, que reúnen la información más relevante para los públicos de primer nivel.

Por último se diseña el sitio dando un lugar de privilegio a las facetas, y un lugar secundario a la clasificación de base, de modo de que los públicos primarios encuentren rápidamente lo que buscan, cuidando que los públicos secundarios tengan en todo momento un camino de navegación hacia las páginas que les interesan.

Los ejemplos de la aplicación de este enfoque de IA son casi interminables. Alcanza con pensar al azar en dos o tres organizaciones de gran tamaño: una de ellas probablemente tenga su sitio organizado según esta propuesta. ¿Será que se trata de un modelo canónico de arquitectura de la información?

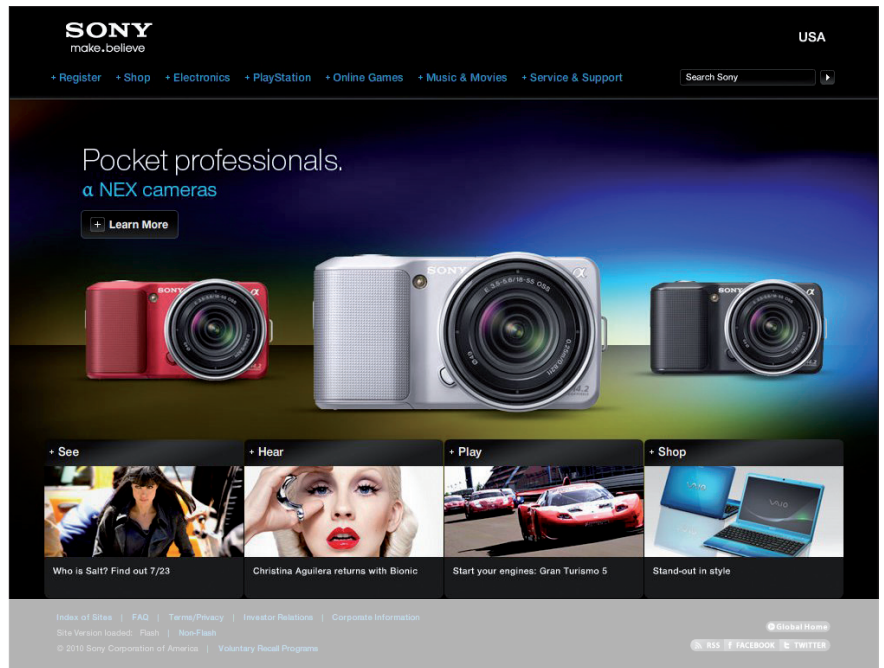


Figura 6. Página principal de Sony, USA, <http://www.sony.com>

5. Notas

1. Taylor, Arlene G. *Introduction to cataloging and classification*, 8th ed. Englewood, Colorado: Libraries Unlimited, 1992.
2. Faceted classification - Wikipedia. http://en.wikipedia.org/wiki/Faceted_navigation
3. Colon significa en inglés coma. El término *Colon classification* surge porque los nombres de las facetas se agregaban separados por comas.
4. En Uruguay, el *Instituto de Seguridad Social* figura en la Constitución de la República como "Banco de Previsión Social" o "BPS". A pesar de que no es un banco ya que ni capta ahorro ni coloca dinero, modificar su sigla para alinearla al nombre que representa solamente el área requiere una reforma constitucional.

6. Bibliografía citada

- Grossman, David A.; Frieder, Ophir. *Information retrieval: algorithms and heuristics*. En particular el capítulo "Retrieval utilities". Springer; 2nd ed., Dec. 2004.
- Hearst, Marti A. *Search user interfaces*. Cambridge University Press, 2009.
- Krug, Steve. *Don't make me think. A common sense approach to web usability*. En especial el apartado sobre pestañas (en inglés *tabs*). New Riders Press; 2nd ed., Aug. 2005.
- Morville, Peter; Rosenfeld, Louis. *Information architecture for the World Wide Web: designing large-scale web sites*. O'Reilly Media; 3rd ed., Nov. 2006.
- Segaran, Toby. *Programming collective intelligence: building smart web 2.0 applications*. En particular los apartados dedicados a las técnicas de clustering y clasificación. O'Reilly Media; 1st ed., Aug. 2007.
- Smiley, David; Pugh, Eric. *Solr 1.4 enterprise search server*. En particular el apartado sobre clustering y facetado de resultados de búsqueda. Packt Publishing, Aug. 2009.
- Daniel Mordecki, Concreta, Montevideo, Uruguay.**
daniel@concreta.com.uy
<http://www.concreta.com.uy>